

What is claimed is:

1. A method comprising:

identifying a plurality of data sources, wherein at least a first data source is more accurate than a second data source;

5 identifying a plurality of overlapping attribute segments to use for comparing the data sources;

calculating a factor as a function of each of the plurality of overlapping attribute segments; and

10 using the factors to update a first group of values in the second data source to reduce bias.

2. The method of claim 1, further comprising:

using the factors to update a second group of values in the second data source to reduce incompleteness.

15

3. The method of claim 2, wherein the first and second groups of values are the same group.

4. The method of claim 2, wherein the first and second groups of values have at

20 least some overlap.

5. The method of claim 2, wherein the first data source is point-of-sale data and the second data source is consumer panel data.

6. The method of claim 2, wherein the first and second sets of values are based on
5 volume.

7. The method of claim 2, wherein the first and second sets of values are based on dollars.

10 8. The method of claim 1, wherein the first data source is point-of-sale data and the second data source is consumer panel data.

9. The method of claim 1, wherein the first data source is manufacturer-supplied shipment data.

15

10. The method of claim 1, wherein the first data source is retailer supplied data.

11. The method of claim 1, wherein the second data source is survey-based data.

20 12. The method of claim 11, wherein the second data source includes causal or promotional data.

13. The method of claim 1, wherein the second data source is sample-based data.

14. The method of claim 13, wherein the second data source includes causal or promotional data.

5

15. The method of claim 1, wherein the first and second sets of values are based on volume.

10 16. The method of claim 1, wherein the first and second sets of values are based on dollars.

17. The method of claim 1, wherein the data sources are compared based on a common time dimension.

15 18. The method of claim 1, wherein the data sources are compared based on a common venue dimension.

19. The method of claim 1, wherein the data sources are compared based on a common product dimension.

20

20. A method comprising:
receiving point-of-sale data and panel data on a periodic basis;

identifying a plurality of product identifiers and a plurality of attributes to analyze;
retrieving and summarizing the point-of-sale data and the panel data by the plurality of
product identifiers, the plurality of attributes, and a plurality of corresponding attribute
segments for a specified time period;

5 calculating a factor for each attribute segment of a particular attribute; and
applying the factors for the particular attribute segment to the panel data to correct
panel bias.

21. The method of claim 20, further comprising:
10 repeating the calculating and applying steps for each attribute segment of the plurality of
attributes.

22. The method of claim 21, wherein the factor for each attribute segment is
calculated by dividing a point-of-sale volume amount by a corresponding panel volume
15 amount.

23. The method of claim 21, wherein the factor for each attribute segment is
applied to the panel data by multiplying each of a plurality of panel data volume amounts by
the factor for the corresponding attribute segment.

20

24. The method of claim 21, further comprising:

prior to said retrieving and summarizing step, preparing a table mapping the product identifiers to the corresponding attribute segments.

25. The method of claim 21, wherein the factors are only calculated for each attribute that is determined to be relevant.

26. The method of claim 21, wherein the factors are only calculated for each attribute segment that is determined to be significant.

27. The method of claim 21, further comprising:
after calculating each attribute segment factor, saving the attribute segment factor in a database.

28. The method of claim 21, further comprising:
applying at least one of the factors to the panel data to correct incompleteness.

29. The method of claim 21, wherein when at least two measures are available for a same factor, a blended factor is calculated and is used to reduce bias in the panel data.

30. The method of claim 29, wherein when calculating the blended factor, a more accurate data source is given a higher relative weight and a less accurate data source is given a lower relative weight.

31. The method of claim 21, wherein the product identifiers are uniform product codes.

5 32. The method of claim 21, wherein the product identifiers are SKU numbers.

33. A method comprising:

receiving point-of-sale data and panel data on a periodic basis;

identifying a plurality of product identifiers and a plurality of attributes to analyze;

10 retrieving and summarizing the point-of-sale data and the panel data by the plurality of product identifiers, the plurality of attributes, and a plurality of corresponding attribute segments for a specified time period;

calculating a factor for each attribute segment of the plurality of attributes; and

applying the factors for each particular attribute segment to the panel data to correct

15 panel bias.

34. The method of claim 33, wherein the factor for each attribute segment is calculated by dividing a point-of-sale volume amount by a corresponding panel volume amount.

20

35. The method of claim 33, wherein attribute segment factors are applied to the panel data by multiplying each of a plurality of panel data volume amounts by the factors appropriate for the corresponding attribute segment.

5 36. The method of claim 33, wherein the factors are only calculated for each attribute that is determined to be relevant.

37. The method of claim 33, wherein the factors are only calculated for each attribute segment that is determined to be significant.

10 38. The method of claim 33, further comprising:
applying at least one of the factors to the panel data to correct incompleteness.

39. The method of claim 33, wherein the product identifiers are uniform product
15 codes.

40. The method of claim 33, wherein the product identifiers are SKU numbers.

41. A method, comprising:
identifying a plurality of product identifiers and a plurality of attributes to analyze for
20 at least two data sources, wherein at least a first data source is more accurate than a second data source;

retrieving and summarizing the first data source and the second data source by the plurality of product identifiers, the plurality of attributes, and a plurality of corresponding attribute segments for a specified time period;

calculating a plurality of factors, wherein one factor is calculated for each attribute
5 segment of the plurality of attributes;
applying the factors to the second data source to reduce bias; and
applying the factors to the second data source to reduce incompleteness.

42. The method of 41, further comprising:
10 determining that additional data sources are available for comparison;
using the additional data sources to calculate additional factors; and
applying the additional factors to the second data source.

43. The method of claim 41, wherein the factor applying steps are repeated
15 multiple times to further correct the second data source.

44. The method of claim 41, wherein the first data source is point-of-sale data and the second data source is consumer panel data.

20 45. The method of claim 41, wherein the product identifiers are uniform product codes.

46. The method of claim 41, wherein the product identifiers are SKU numbers.

47. A system comprising:

one or more servers being operable to store retail data from at least two data sources,
5 store product identifier and attribute categorizations, and store a plurality of factor
calculations;

wherein the at least two data sources includes a first data source that is more accurate
than a second data source; and

wherein one or more of said servers contains business logic that is operable to identify
10 and retrieve a plurality of overlapping attribute segments to use for comparing the at least two
data sources, compare each of the overlapping attribute segments, calculate a factor for each
of the overlapping attribute segments, and use the factors to update a first group of values in
the second data source to reduce bias.

15 48. The system of claim 47, wherein the one or more servers are further operable
to use the factors to update a second group of values in the second data source to reduce
incompleteness.

49. The system of claim 47, wherein the one or more servers are further operable
20 to calculate the factor for each overlapping attribute segment by dividing a first data source
volume amount by a corresponding second data source volume amount.

50. The system of claim 47, wherein the one or more servers are further operable to apply the factor for each attribute segment to the second data source by multiplying each of a plurality of second data source volume amounts by the factor for the corresponding attribute segment.

5

51. The system of claim 47, wherein the one or more servers are further operable to only calculate the factors for each attribute that is determined to be relevant.

52. The system of claim 47, wherein the one or more servers are further operable to only calculate the factors for each attribute segment that is determined to be significant.

10

53. The system of claim 47, wherein the one or more servers are further operable to save each attribute segment factor in a database.

15

54. The system of claim 47, wherein the one or more servers are further operable to apply at least one of the factors to the second data source to correct incompleteness.

55. The system of claim 54, wherein the one or more servers are further operable to calculate a blended factor when at least two measures are available for a same factor, said blended factor being used to reduce bias in the second data source.

20

56. The system of claim 55, wherein the one or more servers are further operable to calculate the blended factor by giving the more accurate data source a higher relative weight and by giving the less accurate data source a lower relative weight.

5 57. The system of claim 47, wherein the one or more servers are further operable to calculate a blended factor when at least two measures are available for a same factor, said blended factor being used to reduce bias in the second data source.

58. The system of claim 57, wherein the one or more servers are further operable
10 to calculate the blended factor by giving the more accurate data source a higher relative weight and by giving the less accurate data source a lower relative weight.

59. The system of claim 47, further comprising:
a number of user stations coupled to the one or more servers over a network, wherein
15 each of the user stations includes a respective one of a number of displays, and is operable to present a graphical user interface that allows a user to administer a plurality of settings used for analyzing and correcting the data sources.

60. The system of claim 47, further comprising:
20 a number of user stations coupled to the one or more servers over a network, wherein each of the user stations includes a respective one of a number of displays, and is operable to

present a graphical user interface that allows a user to view the corrected data in a multi-dimensional format.

61. An apparatus, comprising: a device encoded with logic executable by one or
5 more processors to: identify and retrieve a plurality of overlapping attribute segments to use for comparing at least two data sources, wherein the at least two data sources includes a first data source that is more accurate than a second data source, compare each of the overlapping attribute segments, calculate a factor for each of the overlapping attribute segments, and use the factors to update a first group of values in the second data source to reduce bias.

10

62. The apparatus of claim 61, wherein the device has further logic that is operable to apply all of the factors to the second data source to correct incompleteness.

63. The apparatus of claim 61, wherein the device includes a removable memory
15 device carrying a number of processor executable instructions to define the logic.

64. The apparatus of claim 61, wherein the removable memory device includes a disk.

20 65. The apparatus of claim 61, wherein the device is in the form of one or more parts of a computer network carrying one or more signals encoding the logic.